

# ON THE GENEALOGY OF BRANCHING POPULATIONS AND THEIR DIFFUSION LIMITS

BY OLAV KALLENBERG

---

## Abstract

We consider the evolution of the ancestral structure of a classical branching process in space and its diffusion limit. We also indicate how the conditional structure of the past can be described asymptotically in terms of suitable uniform Brownian trees.

*Keywords:* Branching trees and processes; genealogy; superprocess

2010 Mathematics Subject Classification: Primary 60J68  
Secondary 60J80

---

*Dedicated to my longtime mentor, colleague, and friend Peter Jagers*

The aim of this paper is to outline how the genealogy of certain population processes can be constructed from the simplest birth-and-death processes. We start by considering a classical *binary splitting* branching process  $X = (X_t)$  with *exponential life lengths*, known to be a birth-and-death process with rates  $n\lambda$  and  $n\mu$ . For our purposes, there are two basic cases.

- *Critical splitting process* ( $\lambda = \mu = 1$ ). In this case the mean is constant, yet the process dies out in finite time.
- *Yule process* ( $\lambda = 1, \mu = 0$ ). In such a process there are no deaths: the process keeps growing exponentially.

The former is arguably the most basic of all elementary branching processes. It is also essentially the only case that allows explicit computation, as we now indicate.

**Theorem 1.** *Let  $X = (X_t)$  be a critical binary splitting process with rates 1, and define  $p_t = 1/(1 + t)$ . Then, with  $\mathbb{P}_1\{\cdot\} := \mathbb{P}\{\cdot \mid X_0 = 1\}$  and, for finite  $t > 0$ ,*

- (i)  $\mathbb{P}_1\{X_t > 0\} = p_t$ ;
- (ii)  $\mathbb{P}_1(X_t = n \mid X_t > 0) = p_t(1 - p_t)^{n-1}, n \in \mathbb{N}$ ;
- (iii) *for fixed  $t$ , the ancestors of  $X_t$  at times  $s \in [0, t]$  form a Yule process  $Z^t = (Z_s^t)$  on the time scale*

$$A_s^t = -\log(1 - p_t s), \quad s \in [0, t].$$

Here (i) and (ii) are elementary, and they remain asymptotically valid for certain more general branching mechanisms. Part (iii) is a remarkable and less obvious consequence.

Now let the individuals of our binary splitting process perform independent *Brownian motions* in  $\mathbb{R}^d$ . Then the population evolves as a set of *branching random trees*  $\xi = (\xi_t)$ , in which each  $\xi_t$  is a point process in  $\mathbb{R}^d$  (in all spatial branching processes considered here and below, it is understood that the branching trees originating from different individuals are conditionally independent, given the starting points of those individuals).

Theorem 1 now yields a corresponding statement about such spatial branching processes.

**Corollary 1.** *Let  $\xi = (\xi_t)$  be a set of critical branching random trees in  $\mathbb{R}^d$  with binary splitting at rates 1 and with independent Brownian motions along the branches. Define  $p_t = 1/(1 + t)$  and fix any  $t > 0$ . Then, for  $s \in [0, t]$  and setting  $h = t - s$ ,*

- (i) *the ancestors of  $\xi_t$  at time  $s$  form a  $p_h$ -thinning  $\zeta_s^t$  of  $\xi_s$ ;*
- (ii)  *$\xi_t$  is a sum of conditionally independent clusters of age  $h$ , rooted at the points of  $\zeta_s^t$ ;*
- (iii) *the point processes  $\zeta_s^t$  with  $s \in [0, t]$  combine into a Brownian Yule tree on the time scale*

$$A_s^t = -\log(1 - p_t s), \quad s \in [0, t].$$

Recall that, for any constant  $p \in [0, 1]$ , a  $p$ -thinning of a point process  $\zeta$  is obtained by individual deletion of the points of  $\zeta$ , independently with probability  $1 - p$ . The *clusters* of age  $h$  are independent copies of the original process, each starting from a single individual at time  $s$ . The ancestral processes  $\zeta_s^t$  clearly combine into branching trees on  $[0, t]$ , but their nonhomogeneous Markov property may be less obvious, making them into time-changed Yule processes. The time scale  $A$  determines the splitting rate of the process  $\zeta^t = (\zeta_s^t)$  at time  $s$ .

Given a branching random tree in  $\mathbb{R}^d$ , as above, starting from a *Poisson* process on  $\mathbb{R}^d$  with intensity  $\mu$ , we now perform a *scaling* on three levels:

1. increase the *initial intensity* by a factor  $n$ ;
2. speed up the *branching/death* by a factor  $n$ ;
3. reduce the *weight* of each particle by a factor  $n^{-1}$ .

As  $n \rightarrow \infty$ , we get convergence to a weakly *continuous*, measure-valued, strong *Markov* process  $\xi = (\xi_t)$  starting at  $\mu$ , called a Dawson–Watanabe superprocess (or *DW process* for short). For  $d \geq 2$  and fixed  $t > 0$ ,  $\xi_t$  is an a.s. *singular, diffuse* random measure on  $\mathbb{R}^d$  of Hausdorff dimension 2.

The DW process  $\xi = (\xi_t)$  may be thought of as a randomly evolving diffuse cloud. As such, it seems to defy any simple probabilistic description. It is then highly remarkable that  $\xi$  can be constructed from an underlying discrete structure, composed by clusters of suitable branching trees. Furthermore, just as the process  $\xi$  itself arises in the limit from suitable discrete branching trees  $X$ , the underlying discrete structure of  $\xi$  appears in the limit from the corresponding cluster structure of  $X$ . In particular, the genealogy of  $\xi$  is a limiting version of the genealogy of  $X$ , which leads us to the following counterpart of Corollary 1.

**Theorem 2.** *Let  $\xi = (\xi_t)$  be a DW process in  $\mathbb{R}^d$ . Fix any  $t > 0$ . Then, for  $s \in [0, t)$  and setting  $h = t - s$ ,*

- (i) *the ancestors of  $\xi_t$  at time  $s$  form a Cox process  $\zeta_s^t$  directed by  $h^{-1}\xi_s$ ;*
- (ii)  *$\xi_t$  is a sum of conditionally independent clusters of age  $h$ , rooted at the points of  $\zeta_s^t$ ;*

- (iii) the point processes  $\zeta_s^t$  with  $s \in [0, t)$  combine into a Brownian Yule tree on the time scale  $A_s^t = -\log(1 - t^{-1}s)$ ;
- (iv) as  $s \uparrow t$ ,  $h^{-1}\zeta_s^t \xrightarrow{v} \xi_t$  a.s.

Recall that a point process  $\zeta$  is called a *Cox process directed by  $\xi$* , if  $\zeta$  is conditionally Poisson with intensity measure  $\xi$ . This is the obvious limiting case of the  $p$ -thinnings in Corollary 1, which suggests part (i). The clusters in (ii) are no longer ordinary processes, since their *distributions are unbounded*. However, it still makes sense to consider their conditional distributions, given that they survive until an age  $\varepsilon > 0$ . They correspond to the clusters in the decomposition of infinitely divisible random measures.

The inhomogeneous Markov property in (iii) is again less obvious, and it leads in this case to an *unbounded Yule process* (since  $A_s^t \rightarrow \infty$  as  $s \rightarrow t$ ). Thus, the number of ancestors of age  $h$  tends to  $\infty$  as  $h \rightarrow 0$ . Statement (iv) shows that the original process  $\xi$  can be recovered from those ancestral processes by a law of large numbers.

Our final aim is to consider the *conditional genealogy*, given that the population at time  $t > 0$ , represented by the diffuse random measure  $\xi_t$ , contributes to the points  $x_1, \dots, x_n$ . Such conditional distributions may be described by the *nth order Palm distributions* of the underlying discrete cluster structure, as may be justified by a suitable limit theorem. Indeed, the familiar interpretation of univariate Palm distributions of simple point processes in terms of elementary conditioning can be extended to the present context of diffuse random measures.

Those *nth order Palm distributions* are obtained by disintegration of the corresponding higher order *Campbell measures*, an extension of the *nth order moment measures*. The resulting *nth order Palm kernel* is only defined up to a null set, and we need to invoke a suitable *duality argument* to ensure the existence of a *regular version*. It is the latter version that admits the following remarkable representation in terms of random trees.

The relevant trees are now the *uniform Brownian trees* on  $[0, t]$  of order  $n$ , defined as follows:

- (a) any tree is *rooted* at the origin of  $\mathbb{R}^d$  at time 0;
- (b) there are  $n - 1$  *binary splits*, leading to  $n$  *leaves* at time  $t$ ,
- (c) the splitting times are given by a *binomial process* on  $[0, t]$ ,
- (d) there are independent *Brownian motions* in  $\mathbb{R}^d$  along the branches,
- (e) there are *marks*  $1, \dots, n$  attached to the leaves,
- (f) all branching structures are *equally likely*.

Recall that a (uniform) *nth order binomial process* on  $[0, t]$  is a point process of the form  $\sum_{k \leq n} \delta_{\sigma_k}$ , where the  $\sigma_k$  are i.i.d. and uniformly distributed on  $[0, t]$ . Such trees have been used to give a probabilistic *description* of higher-order moment and Campbell measures.

The following result gives the basic connection between the conditional genealogy of a DW process and the uniform Brownian tree. To simplify the statement, we consider only the contributions from a *single cluster* rooted at 0.

**Theorem 3.** *For a single cluster  $\eta$  of a DW process in  $\mathbb{R}^d$ , and fixed  $t > 0$  and  $x_1, \dots, x_n \in \mathbb{R}^d$ , the following statements agree almost everywhere:*

- (i) the *Palm distribution of the genealogy of  $\eta$* , at points  $x_1, \dots, x_n$  in the support of  $\eta_t$ ,
- (ii) the *conditional distribution of the nth order uniform Brownian tree*, given that the leaves reach  $x_1, \dots, x_n$ .

*To obtain the complete family structure, we need to adorn the tree with some independent clusters, rooted along the branches at a constant rate 2.*

For  $n = 1$ , the conditional tree is just a *Brownian bridge* on  $[0, t]$ , going from 0 to  $x$ . The multivariate statement seems to be much deeper. The general result implies in particular a probabilistic *interpretation* of Campbell and moment measures in terms of Brownian trees.

### Sources and historical remarks

This paper is based on material in Sections 13.1–3 of Kallenberg (2017), which contains some further details and selected proofs. Classical *Bienaymé branching processes* (often referred to as *Galton–Watson processes*) in discrete and continuous time, along with their various extensions, are discussed in many excellent books, including Jagers (1975). General extinction probabilities and associated cluster distributions were obtained by Kolmogorov (1938) and Yaglom (1947), respectively.

Watanabe (1968) studied the binary splitting branching process with independent Brownian motions of particles; Watanabe also derived the diffusion limit, here referred to as a DW process. The corresponding *historical process* was discovered and analysed by Dawson and Perkins (1991), Dynkin (1991), and Le Gall (1991).

Etheridge (2000) provided a recursive construction of a path-valued Markov process to clarify Dynkin's (1991) formulae for the moment measures; in fact, her construction is equivalent to our uniform Brownian tree. The proof of Theorem 3 uses an extension of Le Gall's (1991) *Brownian snake*. No elementary proof is known.

### References

- DAWSON, D. A. AND PERKINS, E. A. (1991). Historical processes. *Mem. Amer. Math. Soc.* **93**, 179pp.
- DYNKIN, E. B. (1991). Branching particle systems and superprocesses. *Ann. Prob.* **19**, 1157–1194.
- ETHERIDGE, A. M. (2000). *An Introduction to Superprocesses* (Uni. Lecture Ser. **20**). American Mathematical Society, Providence, RI.
- JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.
- KALLENBERG, O. (2017). *Random Measures, Theory and Applications*. Springer, Cham.
- KOLMOGOROV, A. N. (1938). Zur Lösung einer biologischen Aufgabe. *Izv. Nauch. Inst. Mat. Mech. Tomskom Gosud. Univ.* **2**, 1–6.
- LE GALL, J.-F. (1991). Brownian excursions, trees and measure-valued branching processes. *Ann. Prob.* **19**, 1399–1439.
- WATANABE, S. (1968). A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.* **8**, 141–167.
- YAGLOM, A. M. (1947). Certain limit theorems of the theory of branching random processes. *Dokl. Acad. Nauk SSSR* **56**, 795–798.

OLAV KALLENBERG, *Auburn University*

Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849-5310, USA.

Email address: kalleoh@auburn.edu